Math 55 Week 5 Review Session

Lillian MacArthur

October 2025

Topics covered: Jordan Normal Form, the characteristic polynomial, the minimal polynomial, category theory, bilinear forms

Jordan Normal Form

The idea behind Jordan Normal Form is to diagonalize a matrix *as much as possible*, even if it is not a diagonalizable matrix. The key step is to choose a basis of vectors which lie in the generalized eigenspace of an eigenvalue, which is the closest we can get to having a basis of eigenvectors.

Definition 0.1. A Jordan block of dimension n and with eigenvalue λ is an nxn matrix of the form

$$\begin{bmatrix} \lambda & 1 & 0 & \dots \\ 0 & \lambda & 1 & \ddots \\ 0 & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda \end{bmatrix}$$

I say a Jordan block with eigenvalue λ is in the generalized eigenspace of λ. Why? The generalized eigenspace of λ is all $\varphi \in Hom(V, V)$ such that $(\varphi - \lambda I)^m = 0$ for some $m \in \mathbb{N}$. If ϕ is a Jordan block with eigenvalue λ, then

$$\varphi - \lambda I = \begin{bmatrix} 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \ddots \\ 0 & \ddots & \ddots & 1 \\ 0 & \dots & 0 & 0 \end{bmatrix},$$

the matrix that is entirely zero, except for its upper diagonal, which is all 1. Many of you will recognize it as a nilpotent matrix, indeed, squaring then cubing and so on this matrix will move the diagonal of ones up and up, until this becomes the zero matrix, $(\varphi - \lambda I)^n = 0$. This is why the Jordan blocks lie in the generalized eigenspaces.

Characteristic and Minimal Polynomials

The characteristic polynomial and minimal polynomial are two algebraic expressions that are given by a linear operator. You may have seen their typical definitions before:

- The typical definition of the characteristic polynomial is $\chi_\phi(x) = det(xI - \phi)$. You might have seen and solved this equation many times if you've taken a linear algebra course before. The roots are the eigenvalues, with multiplicity the multiplicity of their generalized eigenspaces.
- The typical definition of the minimal polynomial is the smallest (lowest degree) equation that an operator solves. A similar method to Problem 8 on your last problem set can be used to show that indeed, any polynomial that an operator φ satisfies must be a multiple of its minimal polynomial. Then, the minimal polynomial contains essential information about the algebraic structure of φ.

If you are given a matrix in Jordan Normal form, it is quite straightforward to find the characteristic and minimal polynomials. First, label the eigenvalues $\lambda_1, \dots \lambda_N$. Then, let n_i be the number of times λ_i appears on the diagonal of the matrix. Lastly, let m_i be the size of the largest Jordan block of λ_i. Then, we define

$$\chi_\varphi = \prod_{i=1}^{N}(x - \lambda_i)^{n_i}, \;\; \mu_\varphi = \prod_{i=1}^{N}(x - \lambda_i)^{m_i},$$

and we call $\chi_\varphi(x)$ the characteristic polynomial, and $\mu_\varphi(x)$ the minimal polynomial.

Why do these satisfy the typical definitions of characteristic and minimal polynomial?

- The determinant of a matrix in Jordan Normal Form is just the product of its diagonal entries. For ϕ in Jordan Normal form, $\phi - xI$ is a matrix in Jordan Normal form, but every diagonal entry λ_i replaced by $\lambda_i - x$. Then, $det(\varphi - xI) = \prod_{i=1}^{N}(x - \lambda_i)^{n_i} = \chi_\varphi$.
- For the operator ϕ to evaluate to 0 under some polynomial μ_ϕ, it is necessary and sufficient to show that each of the Jordan blocks evaluate to 0. Recall that a Jordan block with dimension n and eigenvalue λ satisfies $(\varphi - \lambda I)^n = 0$. Then, $\prod_{i=1}^{N}(\varphi - \lambda_i)^{m_i} = 0$, where m_i is the size of the largest Jordan block of λ_i. It can be proved that this is the smallest polynomial such that φ evaluates to zero.

These polynomials are NOT THE SAME. For example, consider the matrix

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

This is a diagonal matrix, which for the moment we can think of as a matrix with two Jordan blocks, both 1-dimensional of eigenvalue 2. Then, the minimal polynomial of A is $(x - 2)$, while the characteristic polynomial of A is $(x - 2)^2$.

The Cayley-Hamilton theorem says that the minimal polynomial divides the characteristic polynomial. This is obvious from the standpoint of Jordan Normal form, of course $m_i \leq n_i$ for all eigenvalues λ_i. However, from the definition that the minimal polynomial is the lowest degree polynomial that ϕ satisfies, and the characteristic polynomial is $det(\phi - xI)$, this is very nontrivial.

Categories and Functors

Already in Math 55, you have seen very similar ideas come up in the context of different mathematical objects. For example, you can find the kernel of both a map between groups and a map between vector spaces. The language of category theory lets us compare and contrast these similarities, and talk about general mathematical ideas with a larger scope.

Definition 0.2. A category $\mathcal{C}$ is a collection of objects, such that for every pair of objects $A, B \in \mathcal{C}$, there is a collection of morphisms $Mor(A, B)$, such that

- Composition of morphisms $Mor(A, B) \times Mor(B, C) \to Mor(A, C)$ given by $(f, g) \to g \circ f$, is associative.
- For all $A \in \mathcal{C}$, there is an identity morphism $id_A \in Mor(A, A)$, so that given $f \in Mor(A, B)$, $f \circ id_A = id_B \circ f = f$.

This is a lot of intimidating new language. However, it is just talking about ideas that you have already seen before.

- Sets are a category! Any mapping of sets is a morphism between objects. We call this category (set).
- Groups are a category! Group homomorphisms are the morphisms between objects. We call this category (grp). Abelian groups are also their own category, we call it (ab).
- Vector Spaces (over a given field k) are a category! Linear maps are the morphisms between objects. We call this category $(vect_k)$.

In general, most of the categories you will see consist of a common mathematical object, with the morphisms structure-preserving maps. Morphisms then encode how structure transfers between these objects, but how do we describe similarities in structure between categories?

We define a functor to be a map between categories, which sends objects to objects and morphisms to morphisms, in a in a way that preserves identities and composition. There are two types of functors, covariant (think normal functor) and contravariant (think flipping functor).

Definition 0.3. A covariant functor is a map between categories $F : \mathcal{C} \to \mathcal{D}$ which sends every object $X \in \mathcal{C}$ to an object $F(X) \in \mathcal{D}$ and every morphism $f \in Mor_{\mathcal{C}}(X, Y)$ to a morphism $F(f) \in Mor_{\mathcal{D}}(F(X), F(Y))$, such that $F(id_X) = id_{F(X)}$ and $F(g \circ f) = F(g) \circ F(f)$.

Definition 0.4. A contravariant functor is a map between categories $F : \mathcal{C} \to \mathcal{D}$ which sends every object $X \in \mathcal{C}$ to an object $F(X) \in \mathcal{D}$ and every morphism $f \in Mor_{\mathcal{C}}(X, Y)$ to a morphism $F(f) \in Mor_{\mathcal{D}}(F(Y), F(X))$, such that $F(id_X) = id_{F(X)}$ and $F(g \circ f) = F(f) \circ F(g)$.

You have already seen functors before! For example, complexification is a functor $vect_{\mathbb{R}}$ to $vect_{\mathbb{C}}$.

Bilinear Forms

Bilinear forms are linear maps between two vectors and the field of constants over which they are defined. They can be considered as a generalization of the dot product.

Definition 0.5. Given a vector space V over a field k, a bilinear form is a map $b : V \times V \to k$ which is linear in each component.

Since bilinear forms can play many roles and endow the vector space with many different types of structure, mathematicians have many different adjectives they use when describing them:

- If a bilinear form $b(v, w)$ satisfies $b(v, w) = b(w, v)$, it is symmetric.

- If a bilinear form $b(v, w)$ satisfies $b(v, w) = -b(w, v)$, it is skew-symmetric.
- If there is some nonzero $v \in V$ such that $b(v, w) = 0$ for all $w \in V$, we say that b is degenerate.
- If $b(v, w) = 0$ for all $w \in V$ implies that $v = 0$, we say that b is non-degenerate.
- If $b(v, v) \geq 0$, for all $v \in V$, we say that b is positive-definite.
- If b is positive-definite, nondegenerate, and symmetric, we say that b is an inner product.

(If you have some free time, look up the song Positive Definite Nondegenerate Symmetric Bilinear Forms)